Appendix – SUPPLEMENTARY MATERIAL

SI Materials and Methods

Labelling and purification of recombinant MS2 capsids

Recombinant MS2 CP was expressed in *E. coli* and the resulting VLPs purified as previously described (1). Purified VLPs were stored as a precipitate in saturated $(NH_4)_2SO_4$ solution. CP_2 was isolated from resuspended VLPs by dissociation with glacial acetic acid (66% v/v) for 1 h on ice, followed by buffer exchange into 20 mM acetic acid using a NAP5 column (GE Healthcare). Negative stain EM was used to confirm that all capsids were dissociated. Only CP_2 fractions with A_{260}/A_{280} ratio ≤ 0.59 were used for assembly reactions to ensure that the protein was also free of nucleic acids.

Lysines on purified VLPs were labelled with Alexa Fluor 488-SDP ester (AF488, Invitrogen) according to the manufacturer's protocol with minor modifications (2). Briefly, a 100 µl VLP aliquot (9 mg/ml) in 0.1 M sodium carbonate, pH8.4, was mixed with 10 µl of freshly dissolved AF488-SDP ester (10 mg/ml stock in anhydrous DMSO) and the mixture agitated at 4^{o}C for 2 h, before being dialyzed against 100 mM ammonium acetate, pH6.8, at 4°C. The degree of labelling was calculated according to the manufacturer's protocol using the following formula:

$$\text{(moles dye per mole protein)} = A_{494}/(\varepsilon_{488} \times [CP]),$$

where [CP] is the molar concentration of the protein, ε_{488} = 71 000 $M^{-1}cm^{-1}$ is the molar extinction coefficient for AF488, A_{494} is the absorbance at 494 nm for a 1 cm path length.

Labelled VLPs were further purified on a sucrose density gradient (15 - 45 % w/v, 18,000 rpm in a SW40Ti rotor, 16 h, at 4°C) to remove species that may have been denatured during labelling. Labelled capsid fractions were dialysed against 20 mM Tris-HCl, 1 mM EDTA, 50 mM NaCl, pH7.8, overnight at 4°C before being analyzed on SDS-PAGE, and visualized by both fluorescence scan and staining with Instant Blue (Triple Red) (Fig. S1A). The concentrations of coat protein dimer were estimated from A_{280} using a molar extinction coefficient for MS2 coat protein dimer, ε=33240 $M^{-1}cm^{-1}$. For AF488-labelled protein a correction for absorption of the dye at 280 nm was applied (0.11 x A_{494}, *A guide to fluorescent probes and labelling technologies*, 10th Edition, Molecular Probes).

In vitro transcription and purification of subgenomic and non-viral RNAs

The 3’ and 5’ subgenomic RNAs were transcribed from DNA templates (pSMART 2676 and pSMART 2578) linearized with *Hind*III, whilst the internal 928 nt fragment template (iRNA) was produced by PCR amplification of pSMART2676 plasmid as previously described (3). The MS2 full genomic cDNA template was produced by reverse-transcribing wt MS2 phage RNA using Superscript III Reverse Transcriptase and random hexamer oligonucleotide primers (Invitrogen). Primers MS2_F1 and MS2_R1 (Table S3) were utilized to amplify the resulting cDNA using high-fidelity Phusion DNA polymerase (New England Biolabs). The resulting PCR product was used as a template in a second amplification with 5’-phosphorylated primers MS2_F2 and MS2_R2 (Integrated DNA Technologies). The obtained PCR product was agarose gel purified using QIAquick gel extraction kit (Qiagen) and ligated into a pSMART HCAmp vector (Lucigen) following the manufacturer’s protocol. XL1 Blue competent cells (Novagen) were used for transformation with the ligated products and the resulting colonies were PCR-screened and plasmid DNA was isolated from the positive clones, sequenced using primers SL1 and SR2 (Lucigen). *Hpa*I digestion of the resulting plasmid yielded a linearized DNA template that was used for *in vitro* transcriptions.

The template for transcription of NVR1 was produced by cloning a portion (equivalent in size to MS2 genomic RNA) of the open reading frame of *E.coli* RNA polymerase B subunit gene (rpoB). Primers rpoB_F1 and rpoB_R1 (Table S3) were used to amplify the selected region of the gene using Phusion DNA polymerase and genomic DNA extracted from BL21 strain of *E.coli* as a template. The resulting PCR product encompassing nucleotides 18-3587 from the rpoB gene was then amplified using 5’-phosphorylated primers rpoB_F2 and rpoB_R2 (Integrated DNA Technologies) in order to add a T7 promoter sequence to the 5’-end and a *Dra*I restriction site to the 3’-end. Further PCR product purification and cloning into pSMART HCAmp vector were performed as described above for the full genomic MS2 DNA template. *Dra*I digestion of the resulting construct yielded a single linearized DNA product which was used as a template for an *in vitro* transcription.

The non-viral RNA 2 (NVR2) was transcribed *in vitro* from an *Nde*I-linearized pGEM-3Zf(+) vector (Promega, GenBank accession number X65304). Similarly, the non-viral RNA 3 (NVR2) was produced by transcribing a plasmid pTRI Xef, containing the 1.89-kbp elongation factor 1-α gene from *Xenopus laevis* (Ambion).

The DNA template for production of the STNV-C genomic RNA was a gift from Dr Robert Coutts (4). The STNV-C RNA was produced by a run-off transcription of the *Xho*I-linearized STNV-C DNA template.

In vitro transcription was carried out using a T7 RNA transcription kit (MEGAscript, Ambion) following the manufacturer's protocol. RNA was purified using RNeasy mini kit (QIAGEN) following the manufacturer's protocol, except for the fluorescently labelled RNAs. For those samples the RNA-loaded column was washed 3 times with acetone and once with 80%(v/v) ethanol prior to elution with 30 µl of sterile nuclease-free water in order to eliminate contamination with free dye. All RNA samples were routinely examined on denaturing formaldehyde agarose gels to ensure their integrity (Fig S1). Every precaution was taken to avoid RNase contamination and RNA samples were kept as 10 µl aliquots at -80°C to minimise degradation. RNA concentrations were estimated from A_{260} with extinction coefficient (25 ml mg^{-1} cm^{-1}) and the corresponding molar concentrations were calculated using estimated molecular masses (Ambion molecular weight calculator): MS2 genomic RNA = 1148.3 kDa, 5'RNA = 794.3 kDa, 3' RNA = 829.5 kDa, iRNA = 298.5 kDa, NVR1 = 1079.3 kDa, NVR2 = 831 kDa, NVR3 = 607.4 kDa, STNV-C genomic RNA = 397.4 kDa.

Purification of MS2 virions and viral RNA

Wild type MS2 phage was propagated in a permissive *E. coli* C3000 strain and purified as described elsewhere (5). Freshly purified virus was used for RNA extraction using the RNeasy mini kit (QIAGEN) following the manufacturer's protocol. Heat denatured RNA was prepared by rapid heating of 1-2 µM RNA samples in a total volume of 20 µl in RNAse free 10 mM Tris-HCl buffer, pH7.5, to 85°C for 5 min and then snap-cooling on ice and storing it at -80°C. Non-denatured RNA was stored under the same conditions except it was incubated at room temperature for 5 min instead of being heated. Both samples were analysed on a denaturing formaldehyde gel to confirm their integrity after incubation.

STNV coat protein purification

BL21 (DE3) pLysS cells (Promega) harbouring pET22b-STNV1 vector were used to express STNV coat protein following the original protocol (6). Further purification of STNV VLPs was done following this protocol with few modifications. After sonication and DNAse I treatment, the resulting cell lysate was vigorously mixed with chloroform (1:1 v/v) and the emulsion was clarified by centrifugation at 4000 g. The supernatant containing VLPs was

dialyzed against deionised distilled water at 4^{o}C and then incubated for 30 min with polyethyleneimine (PEI), final concentration of 0.05% w/v, pH=9.0. PEI was pelleted for 30 min at 4000g and the VLPs in the supernatant were precipitated with $(NH_4)_2SO_4$, final concentration 1.75 M. The VLP pellet was resuspended in 50 mM Tris-HCl, pH=7.5 and further purified on a sucrose density gradient, 15-45 % (w/v), 20,000 rpm, SW32Ti rotor, 15 h, 4°C. Capsid fractions were dialysed against 10 mM Tris-HCl, pH=7.5, overnight at 4°C. Dissociation of VLPs was performed following the protocol described previously (7). Dissociated STNV VLPs were then loaded onto 1 ml resource S column (GE Healthcare), the column was extensively washed with 50 column volumes of buffer A (50 mM NaCl, 25 mM HEPES-NaOH, pH=8.5) and a linear gradient (0-100% of 20 column volumes) of buffer B (1 M NaCl, 25 mM HEPES-NaOH, pH=8.5) was applied to elute the bound coat protein. Peak fractions were concentrated using Amicon Ultra-4 filter units (Millipore) and 0.5 ml of concentrated STNV protein was injected onto Superdex 200 300 GL column (GE Healthcare) pre-equilibrated and run with buffer C (25 mM HEPES-NaOH, pH=7.5, 100 mM NaCl, 4 mM $CaCl_2$) at 0.5 ml/min flow rate at 4^{o}C. The main peak fraction containing STNV coat protein was analyzed on SDS-PAGE stained with Instant Blue (Triple Red) (Fig. S1B). The concentrations of coat protein were estimated from A_{280} using a molar extinction coefficient for STNV coat protein, ε=11460 $M^{-1}cm^{-1}$. The A_{260}/A_{280} ~0.57 indicated that the STNV coat protein was free of nucleic acids. All the buffers used for STNV coat protein purification were RNAse-free and the columns were pre-treated with 1 M NaOH and extensively washed with DEPC-treated water to minimize the risk of contaminating samples with RNAses.

Fluorescent-labelling of the RNAs

Two end-labelling protocols were developed and routinely used for fluorescent labelling of RNAs. The first relies on the ligation of a 5'-phosphorylated-oligo-dA(25-mer)-3'-AF488 oligonucleotide to the 3’-end of the RNA using T4 RNA ligase (New England Biolabs). The ligation reaction was allowed to proceed for 16 h at 16°C, followed by RNA precipitation with LiCl (8). The RNA pellet was resuspended in 100µl of RNAse-free water and purified using RNeasy columns (QIAGEN), see above.

Labelling of the genomic RNA using the above protocol proved difficult, perhaps due to the presence of secondary and/or tertiary structures. Hence, an alternative approach was adopted in which a reactive amine group was incorporated at the 3’-end,

followed by chemical coupling to AF488-SDP-ester. Genomic RNA was initially modified with addition of 3'-NH_2-ATP (Biolab LSI, Germany) using poly-A polymerase (New England Biolabs) at the 3' end following manufacturer's protocol. The reaction was allowed to proceed for 1 h at 37°C, before the RNA was purified using an RNeasy column and eluted with nuclease-free water. The 3'-end amine modified RNA (1 µM) in a total volume of 100 µl 100 mM RNAse-free sodium borate, pH 8.4, was reacted with 2 mM AF488-SDP ester (dissolved in DMSO) for 3 h at 25°C, then precipitated with LiCl as described above. After resuspension in 100 µl of RNAse-free water the labelled RNA was purified using RNeasy columns following the modified purification protocol for fluorescently labelled RNA. The concentration of fluorescently labelled RNA was calculated using the following formula: C (µg/ml) = 40 x (A_{260} – (0.3 x A_{494})) (Invitrogen).

Initially the 3'-NH_2-ATP approach was also used to label STNV RNA. However, the dye at 3' end proved to be significantly quenched and an additional approach has been devised to confirm the results. In this approach 1 mM 5-(3-aminoallyl)-UTP (Ambion), was added to a standard MEGAScript transcription reaction (7.5 mM of each of the nucleotide triphosphates) to be randomly incorporated at sub-stoichimetric levels. The 5-(3-aminoallyl)-uridine was subsequently labelled using Alexa Fluor 488-SDP ester as described above, resulting in a trace labelled RNA. After accounting for quenching in the case of 3'-end labelled STNV RNA both approaches gave results identical within experimental error.

Oligonucleotide synthesis and purification

5'-end amino-modified TR RNA oligonucleotide: 5'-ACAUGAGGAUUACCCAUGU-3' was chemically synthesized and purified as described previously [Murray, 1995 #735]. The purified material was labelled with AF488-SDP ester following manufacturer's protocol. Briefly, modified RNA (25 µg) was resuspended in 100 µl of 100 mM sodium borate, pH=8.5, and mixed with 10 µl of 10 mM AF488-SDP ester and left incubating for 4 h at room temperature with constant agitation. The reaction was quenched by adding ammonium acetate to a final concentration of 2.5 M, the RNA was isopropanol precipitated and recovered as a pellet by centrifugation as described elsewhere (8). Labelled RNA was separated from unreacted species by anion-exchange chromatography (DNAPac PA100, Dionex Corp.). The TR fractions were analysed on PAGE (10% w/v) and labelled bands visualised with a fluoroimaging scanner using a 488 nm excitation laser.

Negative stain transmission electron microscopy

A 5 µl droplet of the sample was applied to a carbon-coated grid (Agar Scientific) which had been rendered hydrophilic by incubation under UV light for 30 min. The sample was allowed to adsorb for 1 min, blotted with Whatman No.1 filter paper and immediately stained with a 20 µl droplet of 2%(w/v) uranyl acetate (repeated three times), and subsequently blotted against filter paper and allowed to air-dry. Negatively stained grids were examined in either a JEOL 1200EX transmission microscope operating at 80 kV or Tecnai Spirit G_2 (FEI) microscope operating at 120 kV.

Assembly reactions

Concentrated stocks of labelled RNA or coat protein were first diluted into the assembly buffer and left to equilibrate for 30 min at 21°C prior to FCS measurements. Two concentrations of labelled, long RNA were used (0.5 or 1 nM as indicated in Figure legends), whilst the concentration of the labelled TR was 10 nM. 50 nM of labelled CP_2 and 50 nM of unlabelled TR were mixed in the assembly buffer, equilibrated at 21°C for 5 min, allowing formation of the kinetically trapped TR:CP_2 complex. Labelled TR (10 nM) was supplemented with unlabelled TR (40 nM) and 50 nM of unlabelled CP_2 to prepare the kinetically trapped complex with labelled RNA. Unlabelled long RNAs (0.5 nM) were used for assembly with labelled coat protein (100 nM dimer CP). All assembly reactions took place at 21°C and data were recorded over 2 h or more in a total volume of 200 µl in a sterile, nuclease-free chamber (Lab-Tek II Chambered Coverglass, Nunc). Electron microscopy grids were prepared from aliquots of the same reaction taken at different time points to monitor capsid production in parallel and under the same conditions as used for FCS.

STNV assembly was performed under the same conditions as for the MS2 phage, except the STNV RNA concentration in assembly reactions was 10 nM (due to lower affinity of coat for RNA) and a stoichiometric amount of STNV coat protein (600 nM) was added to the RNA.

smFCS data collection and analysis

FCS technique is a correlation-based method which exploits spontaneous fluctuations in fluorescent signal in order to obtain characteristic time scales (e.g. relaxation times or

rates) for molecular processes (9) (10). The fluctuations are due to changes in the number density within the measurement volume and are related to rates of chemical (e. g. conversion of fluorescent substrate to non-fluorescent by quenching) or photophysical reactions (e.g. dynamics of triplet state formation, ~1 μs time scale). In addition, for a small, fixed measurement volume (e.g. ~1 fL focal volume created by microscope objective in the sample), the number density also changes due to molecules diffusing in and out of the detection volume. In this case the characteristic time for which the signal is correlated is related to the average time the molecule takes to transverse the confocal volume by diffusion (diffusion time). Intuitively, a small molecule (e.g. fluorescent dye), which has a large diffusion coefficient, will spend relatively short time in the measured volume and thus the fluorescent signal will only be correlated during this short period (e.g. <100 μs for a free dye molecule). Conversely, a large, labelled macromolecular assembly diffuses slowly through the volume and consequently produces considerably longer correlations in the fluorescent signal (e.g. milliseconds). This provides a way to selectively size labelled molecules in the presence of other unlabelled species.

In order to obtain the characteristic times the signal fluctuations are correlated, usually by a specialised hardware correlator, and an autocorrelation function (CF, see Fig. S2A for typical examples) is obtained. Due to signal to noise limitations it usually takes 10 to 30 s to collect enough photons to produce single reliable CF. A typical CF decays from the initial plateau, which is inversely related to the average concentration, with characteristic times representing the above mentioned processes, diffusion often being dominant for commercial dyes (e.g. AF488) with suppressed triplet state formation. The relevant time scales are characterized by corresponding correlation times. A correlation time corresponding to diffusion (sometimes also called diffusion time, T_d) is approximated by the midpoint of the steepest part of CF in Fig. S2A and rigorously extracted by fitting a model CF to the experimental data. In addition to the dominant diffusion component, a labelled macromolecule may also exhibit triplet state dynamics (Fig S2A) and sometimes also quenching dynamics (Fig. S2D). These faster photophysics-related CF components are simply accounted for as exponential multiplicative terms, distinct from the dominant diffusion contribution (9) (10). While these contributions are of no interest in this study in other cases they provide valuable information, e.g. rates of conformational changes or chemical reactions (11).

FCS measurements were performed on a custom-built smFCS confocal setup (11). The excitation laser (Sapphire CW blue laser, 488 nm, Coherent, USA) power was set to 65 µW. The immersion oil objective (63x magnification, numerical aperture of 1.4) was used together with low autofluorescence immersion oil (refractive index 1.515, type DF, Cargille Laboratories, USA). The focus position was adjusted to 20 µm from the cover slip inner surface and precisely maintained by a piezoelectric feedback loop (Piezosystems Jena, Germany). The photon count was recorded and analyzed by an ALV-5000 multiple tau digital correlator (ALV-GmbH) used in a single channel mode. Multiple runs of up to 1000 auto-correlation functions with acquisition scan time of 30 s each were recorded for various samples using ALV-correlator software (ALV-5000/E/EPP, ver.3.0). Calibration of the confocal volume was performed by measuring the diffusion time of AF488-SDP dye (0.25 nM in assembly buffer) before and after each data collection run and the calibration was used to convert the apparent diffusion times to hydrodynamic radii using the estimated hydrodynamic radius of the dye ~0.7 nm in the assembly buffer (21°C, η=1 cP, corresponding to D = 300 $\mu m^2/s$ in water at 21°C (12)) and the Stokes-Einstein equation. FCS data were analyzed by non-linear least-squares fitting with a single-component diffusion model autocorrelation function corrected for the triplet state (13) using Matlab (ver 7.11, MathWorks). An additional quenching component was used for fitting the data for STNV 3'-end labelled RNA. The diffusion time values were used for calculating apparent hydrodynamic radii (R_h) that were plotted as a function of assembly time and further analysed in OriginPro ver. 8 (OriginLab Corporation). Assembly time-dependent $R_h(t)$ data were smoothed applying a fast Fourier transform filter (FFT) algorithm with a 5 or 10 point window. Apparent rates of capsid assembly were obtained by single exponential approximation of the original $R_h(t)$ data.

Establishing FCS assay of MS2 assembly

Labelled CP_2 (Fig. S1A) at concentrations 200 µM remained dimeric in assembly buffer for at least 12 h and had an apparent diffusion time of ~0.4 ms (Fig. S2A). This corresponds to a hydrodynamic radius, R_h ~ 2.5 nm, whilst that for dye-labelled capsids (Fig. S2A, red) was R_h ~12 nm. These values agree well with previous measurements by ensemble techniques (3)(Table S1) and electron microscopy of labelled capsids (Fig. S2A, inset i).

As a test of the ability of smFCS to monitor assembly, we added unlabelled TR RNA to the CP_2 at saturating concentrations. Under these conditions this complex is kinetically trapped because all the protein is A/B-like and there are no C/C-like RNA-free CP dimers,

needed to build a *T*=3 shell (14-16) (Fig. 1). Its CF (Fig. S1A) overlaps that of CP_2, as expected from their similar shape (Table S1). In addition, complexes of labelled TR and unlabelled CP_2 had similar R_h values (Table S1). However, when a 100-fold molar excess of unlabelled CP_2 was added to either form of the complex rapid assembly into *T*=3 capsids occurred. After incubation of these mixtures for 2 h their CFs (Fig. S2A, green) were coincident with that of the control *T*=3 capsids as also shown by negative stain electron microscopy (Fig. S2A inset). Electron microscopy (EM) confirmed that assembly of *T*=3 capsids had occurred also with the larger unlabeled genomic (gRNA) and sub-genomic RNAs (Fig. S2B). These results are consistent with the TR-induced coat protein dimer switching model of assembly, demonstrating that smFCS accurately reports on capsid assembly.

Viral and non-viral RNAs were labelled and demonstrated to have the expected sizes by gel electrophoresis (Fig. S1C-D). The corresponding CFs are shown in Fig. S2C-D.

Supplementary Figures and Tables

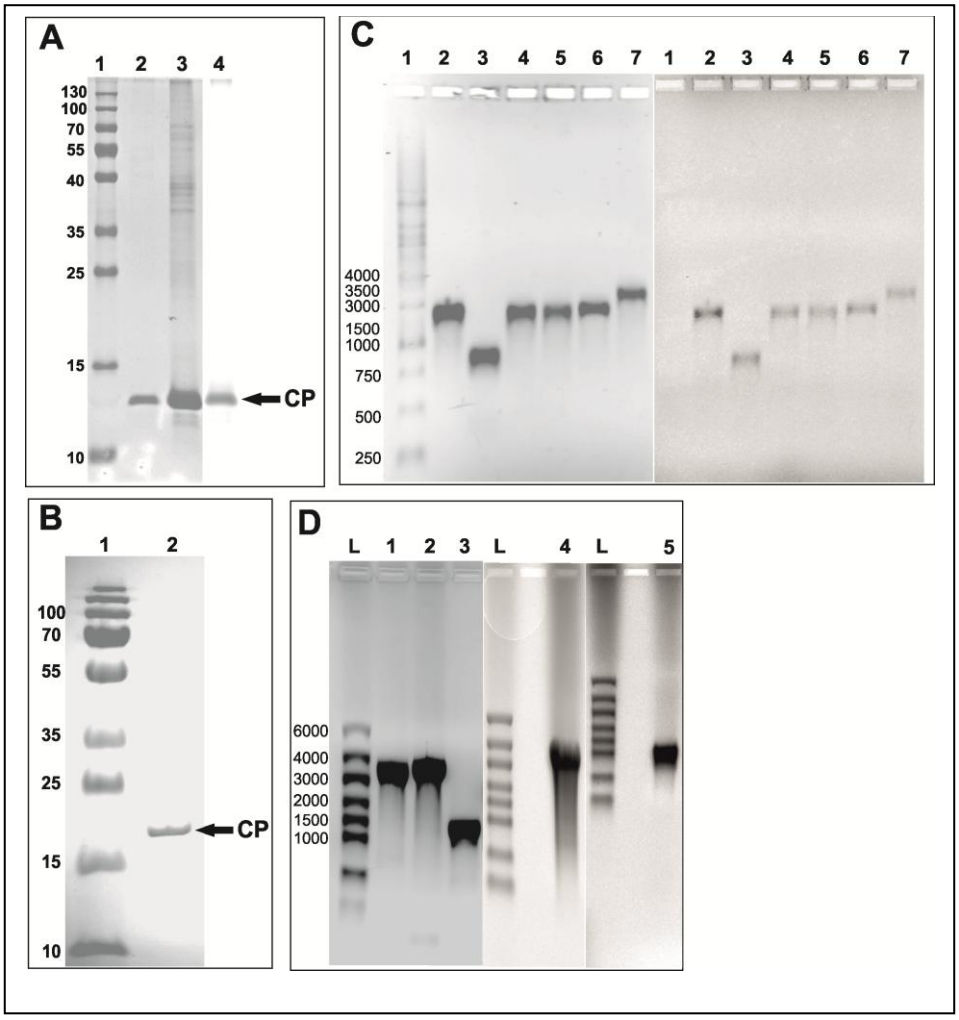


Fig S1: Labelling of capsid proteins and RNA.

(A) SDS-PAGE of AF488-labelled MS2 VLPs purified by a sucrose density gradient centrifugation and stained with Instant blue *lane 1* – pre-stained molecular weight protein marker (Fermentas), sizes indicated in kDa, *lane 2* – purified labelled VLPs; *lane 3* – unlabelled MS2 VLPs; *lane 4* – labelled VLPs as in *lane 2* but visualized by fluorescence. Arrow indicates the band corresponding to the coat protein monomer (MW~13.7 kDa). **(B)** SDS-PAGE of AF488-labelled and purified STNV coat protein. *Lane 1* – molecular weight protein marker (Fermentas), sizes indicated in kDa, *lane 2* – purified labelled coat (stained with Instant blue). **(C)** *In vitro* produced RNAs, labelled at their 3' ends with AF488 were separated on a denaturing 1% (w/v) agarose gel and visualised by staining with SYBR gold (left) or fluorescence (right). *Lane 1* – molecular weight DNA marker (Fermentas), sizes shown in nt. *Lane 2* – 3'RNA, *lane 3* – iRNA, *lane 4* – 5'RNA, *lane 5* – NVR3 RNA, *lane 6* – NVR2 RNA, *lane 7* –MS2 genomic RNA transcript. **(D)** *In vitro* produced RNAs were labelled at their 3' ends with AF488 and separated on a denaturing 1% (w/v) agarose gel and visualised with ethidium bromide (left panel lanes 1-3) or fluorescence (lanes 4 and 5). Lanes 1, 2, 4 - NVR1, lanes 3 and 5 - STNV RNA.

Table S1: Comparison of smFCS hydrodynamic radii to previously published values.

	FCS	AUC*	IMS-MS#
Capsids	12.2±0.9 nm	11.74 nm	N/A
Coat protein (CP_2)	2.5±0.3 nm	2.8 nm	2.48 nm
TR	1.4±0.1 nm	1.21 nm	N/A
CP_2-labelled TR	2.5±0.2 nm	N/A	2.53 nm
Labelled CP_2-TR	2.5±0.6 nm	N/A	2.53 nm

*analytical ultracentrifugation (AUC) (2)

#ion mobility mass spectrometry (IMS-MS) experiments (17)

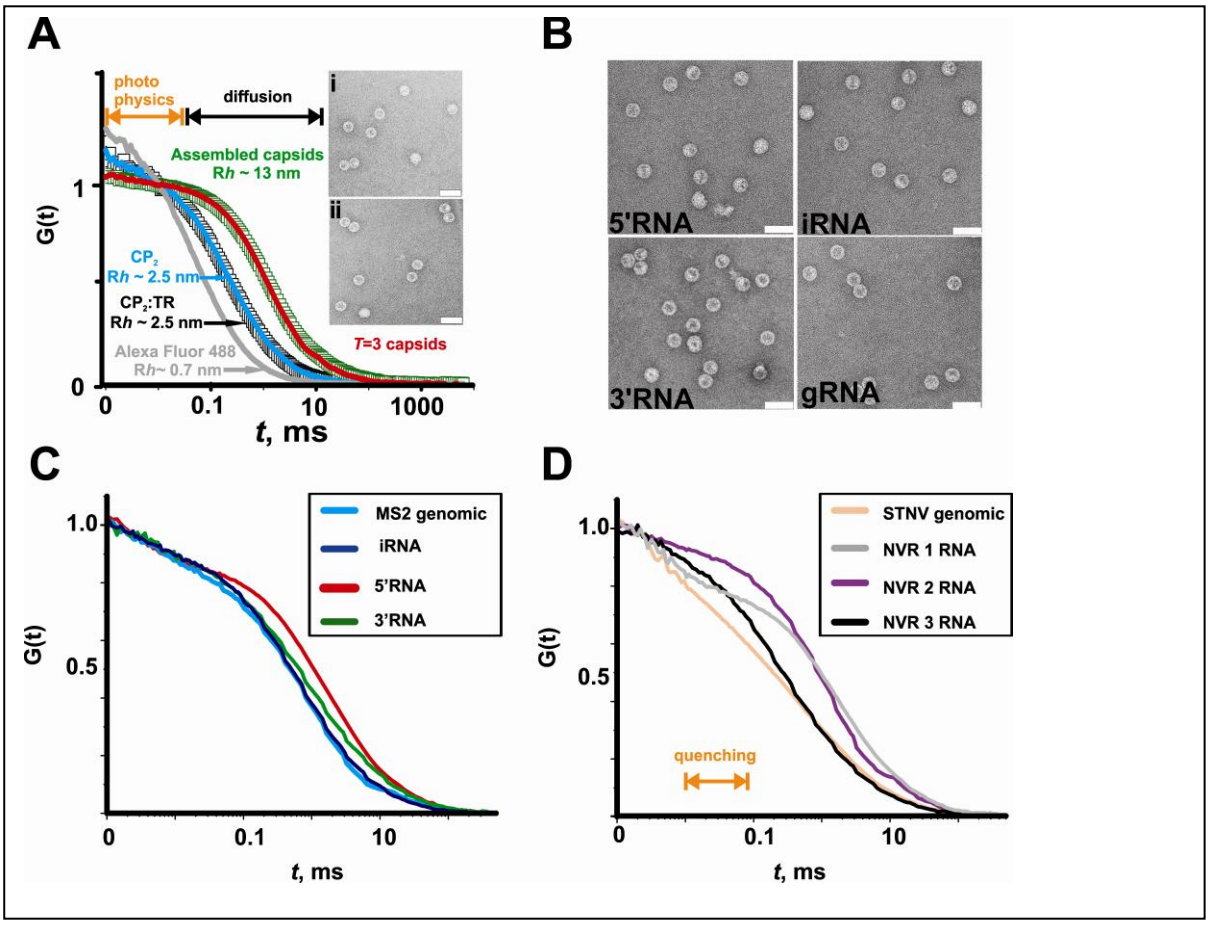


Fig. S2: Validation of smFCS assays of virus assembly.

(A) Examples of fluorescence correlation functions for the AlexaFluor 488 dye (gray), trace-labelled and purified capsids (red line), disassembled coat dimer (blue line), MS2 coat dimer-TR complex (black squares) and capsids re-assembled with 200 nM trace-labelled coat and 10 nM unlabelled TR RNA (green squares) in assembly buffer.

(B) Assembly products with labelled MS2 coat and genomic RNAs examined by negative stain EM. Bar = 50 nm.

(C) Averaged autocorrelation functions (10 x 30 s scans) of labelled MS2 RNAs collected at 1 nM concentration (cyan –genomic RNA, dark blue -iRNA, red – 5'RNA, green – 3'RNA)

(D) Averaged autocorrelation functions (10 x 30 s scans) of labelled RNAs: orange – STNV, gray – NVR1, purple - NVR2, black – NVR3. Orange scale bar indicates the time scale of the additional quenching component that was observed for 3'-end labelled STNV RNA.

Table S2: Average hydrodynamic radii for RNAs as determined by FCS.

RNA	R_h, nm*	Length, nt
MS2 genomic	12.3±0.6 nm	3569
STNV (strain C) genomic	12.3±1.0 nm	1221
5' MS2 transcript	11.9±1.3 nm	2469
3' MS2 transcript	13.8±1.3 nm	2578
iRNA MS2 transcript	8.7±0.8 nm	928
NVR1 RNA	18.3±3.2. nm	3556
NVR2 RNA	12.8±0.1.8 nm	2509
NVR3 RNA	7.9±0.1.8 nm	1890

*Standard deviations represent half widths at half maxima of corresponding R_h distributions (see Fig. S3A for further details).

Table S3: Oligonucleotide primers used for the PCR reactions.

T7 promoter sequences are underlined and the first nucleotide (**G**) incorporated into RNA transcript is shown in bold. The restriction sites utilized for linearizing DNA templates are *italicized.*

Primer name	Primer sequence and restriction enzyme site used
RpoB_F1:	CGAGAAAAAACGTATTCGTAAGGATTTTGG
RpoB_R1	TAATTTCTGCTTCTTTCGCACCGTC
RpoB_F2	TAATACGACTCACTATA**G**GGAGAAAAAACGTATTCGTAAGGATTTTG
RpoB_R2:	***TTTAAA***GCACCGTCGAACACCGGCGTTGCGA (*Dra*I)
MS2_F1	GGGTGGGACCCCTTTCGGGGTCCTGCTCAACTTCCTGTCGAGCTAATGCCATT
MS2_R1	GTAACTAGCCAAGCAGCTAGTTACCAAATCGGGAGAATCCCGGGTCCTCTCTTT
MS2_R2	***GTTAAC***TGGGTGGTAACTAGCCAAGCAGCTAG (*Hpa*I)
MS2_F2	TAATACGACTCACTATA**G**GGTGGGACCCCTTTCGG

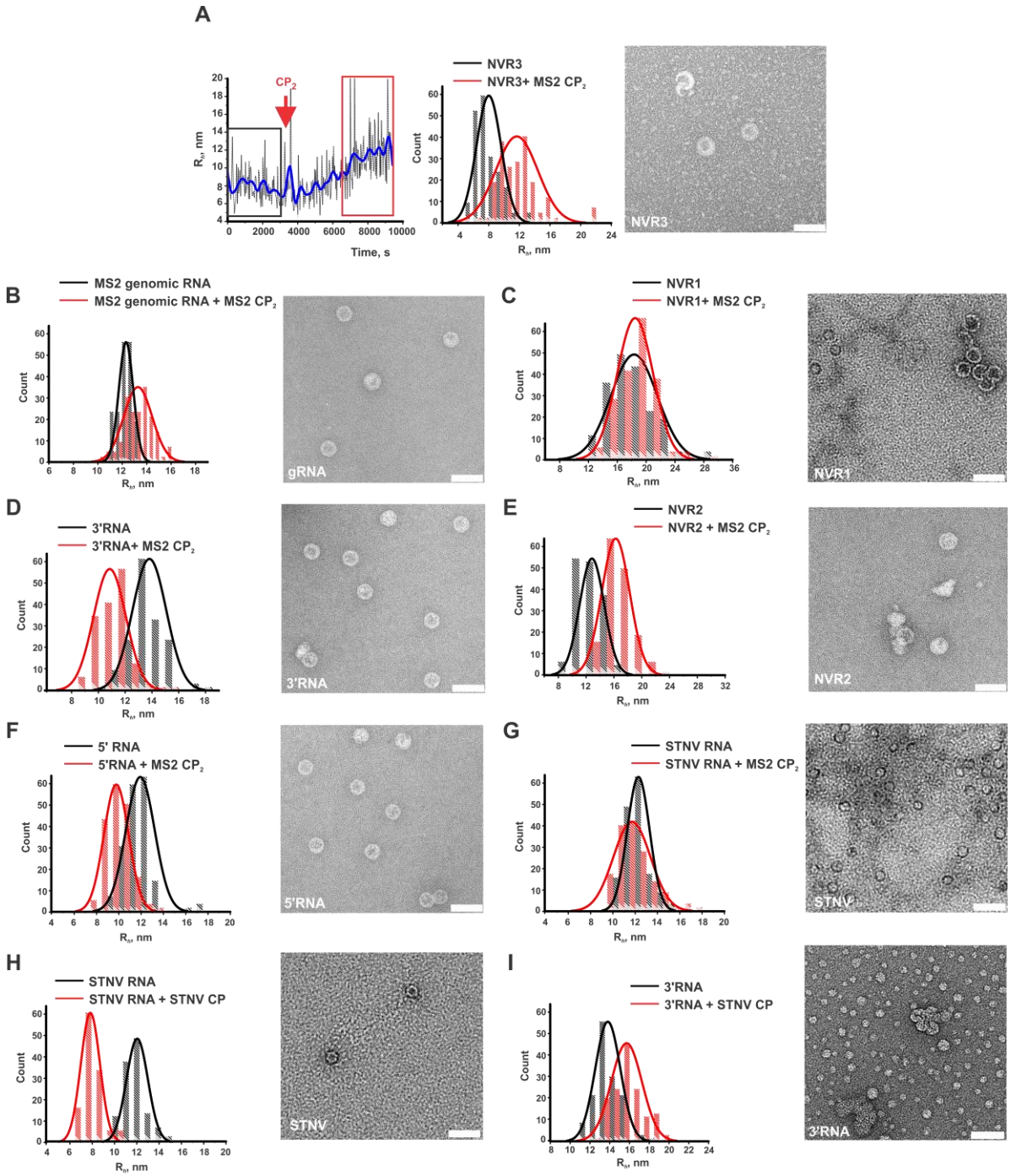


Fig. S3: (A) Left panel shows apparent $R_h(t)$ measured for labelled NVR3 RNA (1 nM) prior to assembly (gray box) and after addition of 200 nM MS2 CP_2 (marked by red arrow). Middle panel: RNA size distributions (histograms) compiled from the data in the gray box in the left panel (100 data points, gray) and red box (last 100 data points, red), which represent the starting RNA conformational ensemble and assembly products, respectively. Right panel: Negative stain EM of assembly reaction products. Bar = 50 nm. **(B-I)** Left panels: Size distributions for various RNAs (as indicated in the panel), computed as described in panel A. Right panels: Negative stain EMs of assembly reaction products. Bar = 50 nm.

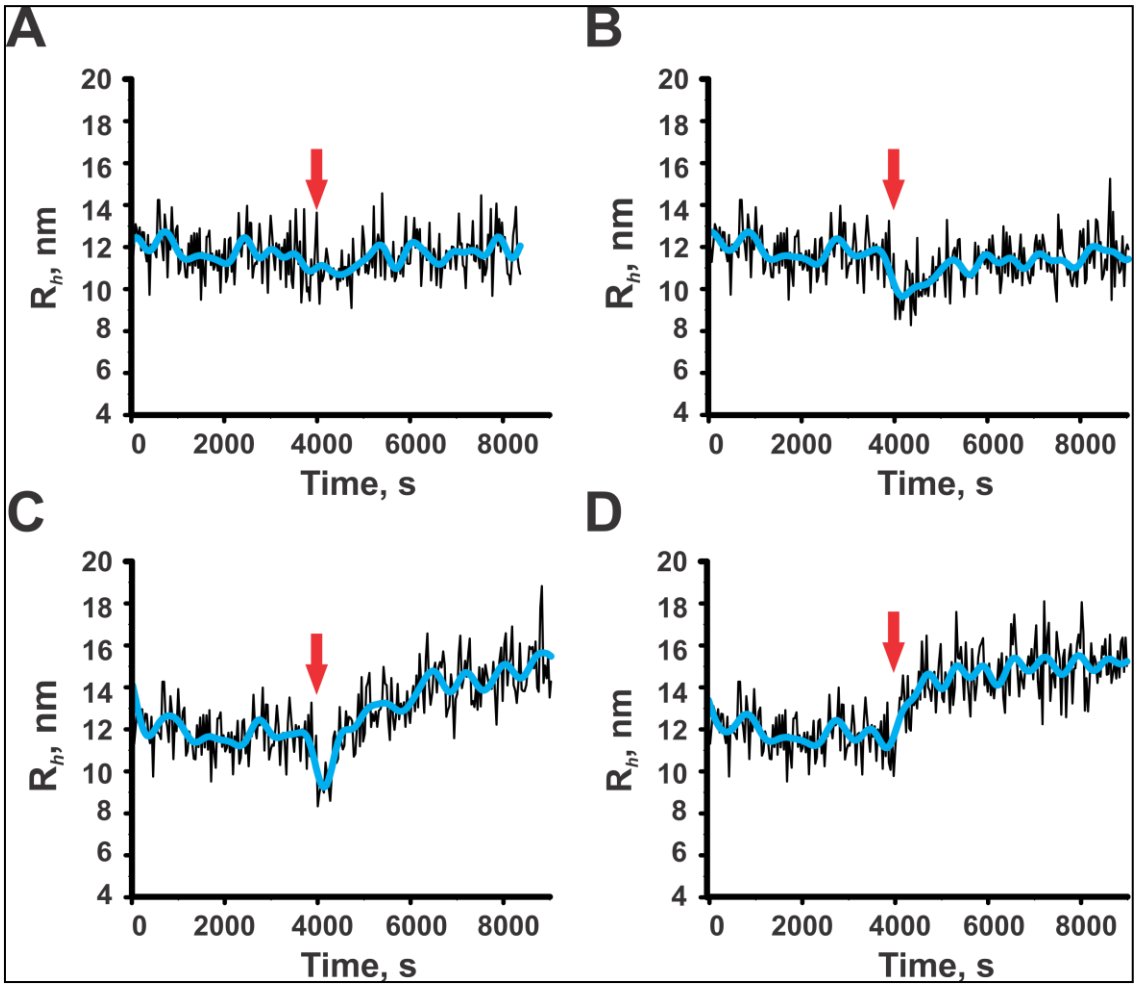
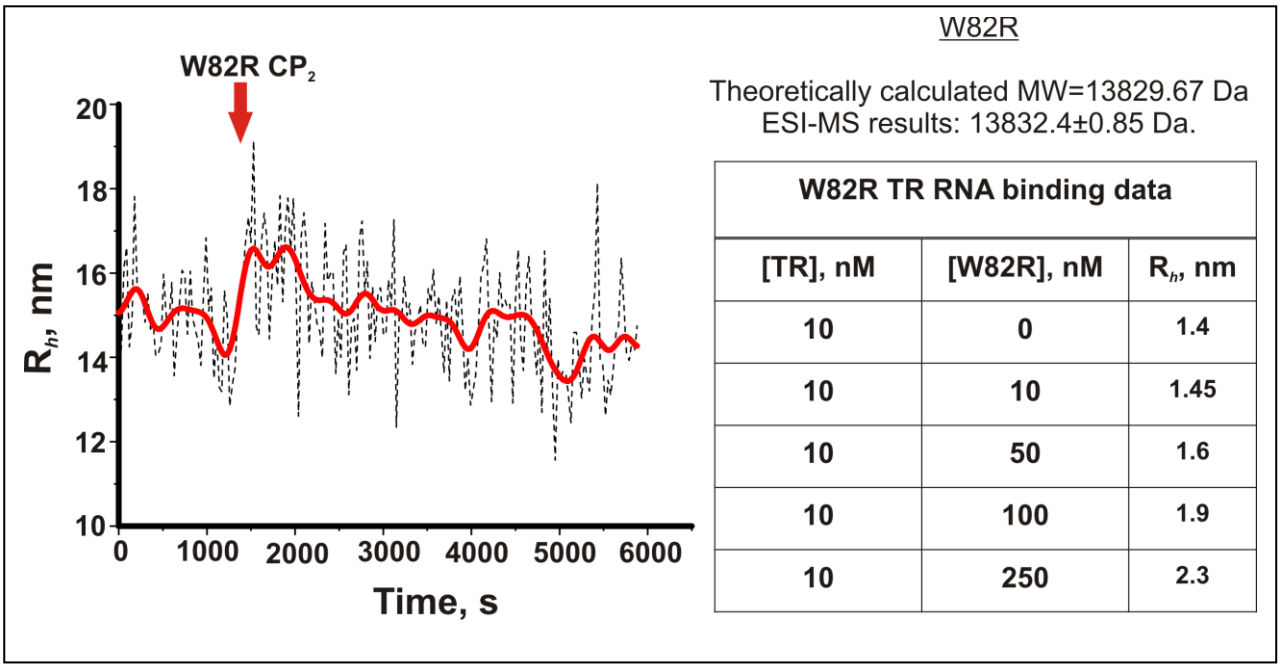


Fig. S4: Dependence of the initial RNA collapse on the coat protein concentration, demonstrated by R_h(t) changes: **(A)** 25 nM, **(B)** 50 nM, **(C)** 100 nM and **(D)** 200 nM [CP_2]. The red arrows indicate when CP_2 was added. At [CP_2]>100 nM the amplitude of the collapse becomes progressively masked by faster second stage kinetics.



W82R

Theoretically calculated MW=13829.67 Da
ESI-MS results: 13832.4±0.85 Da.

W82R TR RNA binding data		
[TR], nM	[W82R], nM	R_h, nm
10	0	1.4
10	10	1.45
10	50	1.6
10	100	1.9
10	250	2.3

Fig. S5: Left panel: Apparent size of labelled MS2 gRNA (1 nM) monitored prior and after the addition (red arrow) of unlabelled CP_2 mutant W82R (200 nM). Right: W82R mutant coat protein exhibits correct mass by mass spectrometry and binds to the labelled TR as shown by the increase of the apparent R_h upon protein addition.

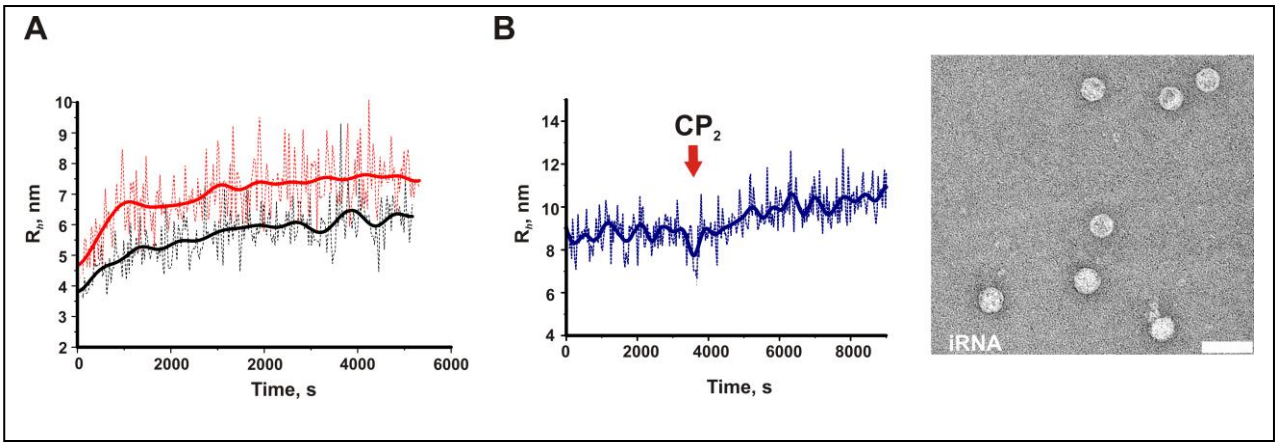


Fig. S6: (A) Assembly of labelled coat (100 nM total [CP_2]) with native (red) and heat denatured (black) virion-derived RNA (1 nM). FFT smoothed data are shown as thick lines. **(B)** Left panel: Assembly of 200 nM coat protein with labelled iRNA. Coat added at the point indicated by red arrow. Right panel: Negative stain EM of the assembly products. Bar = 50 nm.

References

1. Mastico RA, Talbot SJ, Stockley PG (1993) Multiple presentation of foreign peptides on the surface of an RNA-free spherical bacteriophage capsid. *J GenVirol* 74: 541-548.
2. Hermanson GT, Hermanson GT (1996) in *Bioconjugate techniques* (Academic Press, San Diego), p. 785.
3. Rolfsson O, Toropova K, Ranson NA, Stockley PG (2010) Mutually-induced Conformational Switching of RNA and Coat Protein Underpins Efficient Assembly of a Viral Capsid. *J Mol Biol* 401: 309-322.
4. Bringloe DH, et al. (1998) The nucleotide sequence of satellite tobacco necrosis virus strain C and helper-assisted replication of wild-type and mutant clones of the virus. *J Gen Virol* 79: 1539-1546.
5. Strauss JH, Sinsheimer RL (1963) Purification and Properties of Bacteriophage Ms2 and of Its Ribonucleic Acid. *J Mol Biol* 7: 43-&.
6. Lane SW, et al. (2011) Construction and Crystal Structure of Recombinant STNV Capsids. *J Mol Biol* 413: 41-50.
7. Bunka DHJ, et al. (2011) Degenerate RNA Packaging Signals in the Genome of Satellite Tobacco Necrosis Virus: Implications for the Assembly of a T=1 Capsid. *J Mol Biol* 413: 51-65.
8. Sambrook JF ed. (2006) *The Condensed Protocols From Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Press, New York).
9. Gell C, Brockwell D, Smith A (2006) *Handbook of Single Molecule Fluorescence Spectroscopy* (Oxford University Press, Oxford).
10. Rigler R, Elson ES eds. (2001) *Fluorescence Correlation Spectroscopy* (Springer, Berlin).
11. Gell C, et al. (2008) Single-Molecule Fluorescence Resonance Energy Transfer Assays Reveal Heterogeneous Folding Ensembles in a Simple RNA Stem-Loop. *J Mol Biol* 384: 264-278.
12. Doeven MK, et al. (2005) Distribution, lateral mobility and function of membrane proteins incorporated into giant unilamellar vesicles. *Biophys J* 88: 1134-1142.
13. Meseth U, Wohland T, Rigler R, Vogel H (1999) Resolution of fluorescence correlation measurements. *Biophys J* 76: 1619-1631.
14. Stockley PG, et al. (2007) A simple, RNA-mediated allosteric switch controls the pathway to formation of a T=3 viral capsid. *J Mol Biol* 369: 541-552.
15. Basnak G, et al. (2010) Viral Genomic Single-Stranded RNA Directs the Pathway Toward a T=3 Capsid. *J Mol Biol* 395: 924-936.
16. Morton VL, et al. (2010) The Impact of Viral RNA on Assembly Pathway Selection. *J Mol Biol* 401: 298-308.
17. Knapman TW, et al. (2010) Determining the topology of virus assembly intermediates using ion mobility spectrometry-mass spectrometry. *Rapid Commun Mass Spectrom* 24: 3033-3042.